CENTRE for ECONOMIC PERFORMANCE

CEP Discussion Paper No 1175

November 2012

The Enduring Impact of Childhood Experience on Mental Health: Evidence Using Instrumented Co-Twin Data

Rachel Berner Shalem, Francesca Cornaglia and Jan-Emmanuel De Neve

THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

Abstract
The question of whether there is a lasting effect of childhood experience on mental health has eluded causal measurement. We draw upon identical twin data and econometric instrumentation to provide an unbiased answer. We find that 55% of a one standard deviation change in mental health due to idiosyncratic experience at age 9 will still be present three years later. Extending the analysis, we find such persistence to vary with age at impact, gender, and mental health sub-categories. This investigation allows us to get a grasp on the degree to which childhood events influence health and socio-economic outcomes by way of their lagged effect on subsequent mental health. A better understanding of the evolution of mental health also helps identifying when mental health issues can be most effectively treated.

Keywords: mental health, childhood experience, twin study, instrumental variable analysis
JEL Classifications: I10

This paper was produced as part of the Centre's Wellbeing Programme. The Centre for Economic Performance is financed by the Economic and Social Research Council.

Acknowledgements
We gratefully acknowledge the ongoing contribution of the parents and children in the TEDS. For helpful comments and discussion we thank Robert Plomin, Richard Layard, Steve Pischke, Martin Knapp, Steve Machin, Robert Goodman, John Van Reenen, Guy Michaels, Mirko Draca, Sandra McNally, Johannes Spinnewijn, Dan Benjamin, and Amy Challen. We also thank the CEP annual conference participants. For outstanding research assistance we thank Cody Xuereb. All remaining errors are our own. TEDS is supported by a program grant (G0500079) from the UK Medical Research Council; our work on environments and academic achievement is also supported by grants from the US National Institutes of Health (HD44454 and HD46167). Financial support from the Department for Work and Pensions, US National Institute on Aging (Grant R01AG040640), Esmee Fairbairn Foundation, and the Economic & Social Research Council is also gratefully acknowledged.

Rachel Berner Shalem holds an Honors Masters degree in economics from the Hebrew University of Jerusalem. Francesca Cornaglia is a Research Associate with the Centre for Economic Performance, London School of Economics and Political Science. She is also a Lecturer in Economics at Queen Mary University of London. Jan-Emmanuel De Neve is a Research Associate at the Centre for Economic Performance, London School of Economics. He is also Lecturer (Assistant Professor) in Political Economy and Behavioural Science at University College London.

Published by
Centre for Economic Performance
London School of Economics and Political Science
Houghton Street
London WC2A 2AE

Introduction

Mental health is of critical importance to individual well-being, social cohesion, and economic life (1-3) and research has found effects of childhood mental health on physical health and longevity (4-9) as well as education and labour market outcomes (9-10). The mechanism through which mental health impacts later life events, however, remains elusive. Genetic endowments provide some explanatory power for differences in mental health (11-16) but the impact of early environmental experience on later mental health and, in turn, life events has so far eluded causal measurement. Although survey studies have previously shown evidence for an association between childhood experience and adult mental health (17-22), in this study we draw upon identical twins data and econometric instrumentation techniques to provide an unbiased answer to this important question. Evidence for the persistence of mental health experience would bolster the case for treatment approaches and early intervention in the case of childhood experience. If no unbiased evidence can be obtained then the focus should be on preventing negative shocks to mental health instead of their treatment.

In order to better understand the idea of mental health continuity, we illustrate with an example of physical health. Consider four different physical conditions, each of them characterized by a different level of continuity over a period if no treatment were provided: (a) A cold generally goes away within a week whether treated or not: its continuity is zero. (b) A permanent physical impairment such as an amputation has a continuity of one. (c) Degenerative diseases will get worse over time, that is their continuity will be larger than one. (d) A multiple fracture may never completely heal, showing continuity smaller than one, but larger than zero. In this paper we try to understand to which of these four groups idiosyncratic childhood experience belongs and measure its effect on later mental health.

Using the Twins' Early Development Study—a large-scale UK twin pairs data set (23)—we study the persistence of child mental health by performing a comparison of identical twins in order to account for genetic confounding and shared experience. Throughout the analysis we consider a well-known mental health scale—the Strength and Difficulties Questionnaire or SDQ (24-25)—and also evaluate its five specific components: emotional symptoms, conduct disorder, hyperactivity, peer relationship problems, and prosocial behaviour (the SDQ questionnaire and descriptive statistics are provided in figures S1-S2 and tables S1-S3 in the Supplements Appendix). We also take advantage of the availability of a child self-report, a parent report, and a teacher report that evaluate child mental health independently. This unique feature of the data allows us to instrument the analysis, reduce measurement error bias and further limit endogeneity concerns (for methodological proofs see Methods and the Supplements Appendix). For the instrumented co-twin analysis we focus in on a three year timeframe—with study subjects aging from nine to twelve years—for which we have access to all three independent reports on child mental health. To gain a broader sense of mental health continuity we extend our study to as early as 4 years of age though with less measurement accuracy.

Methods

Participants

The Twins' Early Development Study (or TEDS) is a longitudinal study of twins from early childhood through adolescence. TEDS originally enrolled about 15,000 twin pairs born in the UK between 1994 and 1996 to allow for the multivariate investigation of language, cognition and behaviour problems. For a detailed overview of the TEDS cohort and its selection criteria please see the most recent review by Oliver and Plomin (23).

Measures

Child mental health was assessed using the screening instrument called the Strengths and Difficulties Questionnaire (SDQ). The SDQ is a well-validated 25-item questionnaire (24-25). Items are scored 'not true', 'somewhat true', or 'certainly true' and used to construct five subscales of five items each (emotional problems, conduct problems, hyperactivity-inattention, peer problems, and prosocial behaviour). The subscales are also added up to produce a total SDQ score for each informant (See Table S1 in the Supplements Appendix for full item list). The twins' parent (mostly the mother) filled in an SDQ questionnaire starting at age 4, the teacher starting at age 7 and the child filled in a self-reported SDQ questionnaire starting at age 9. Table S3 in the Supplements Appendix summarizes the SDQ averages for our TEDS sample, and compares them with the UK population SDQ averages. The table shows that the TEDS sample is characterized by slightly lower mental health problems (and slightly more prosocial behaviour) than the general population. Figure S2 in the Supplements Appendix also shows that the mean SDQ levels reported by different informants are very different. While the average level of total SDQ according to the parents is 6.5, the average level according to the teachers is 4.7, and as high as 10.7 according to the child self-report measure. The discrepancy between informants is a well-known issue in the psychological literature (see Reyes and Kazdin (26) for a review). The main method the literature uses to investigate these discrepancies is to analyse the correlations between different life factors and the SDQ reports of each informant in different data sets. The differences between the correlations teach us something about the nature of the informant, and possibly about the bias in his/her report. For example, Collishaw *et al.* (27) conduct this analysis with the SDQ measure that we use. In their paper, the authors use the 1999 British Child and Adolescent Mental Health Survey (n = 4,525, ages 11–15 years), where teachers, parents and children completed the SDQ. They show that teachers' reports are more correlated with the child

being male, with the existence of learning difficulties, and with the social economic status of the child's family and neighbourhood. The authors' interpretation of these results is that the teacher is biased by what is considered a risk group (males, poor families), and by the child's behaviour in the class (how well s/he learns). The parent's report is more correlated with the child's poor general health, child's neurodevelopmental disorders and with family variables like the caretaker's distress (usually the reporting parent) and the family's socio-economic status. The authors' interpretation for this is that the parent is biased by other problems that the child has, and also by his/her own mental condition. Since the caretaker's mental condition is also correlated with the socio-economic well-being of the family, it is likely to be correlated with those variables as well. Overall, the literature on this is not conclusive about the reasons why these discrepancies occur, and which of the informants is more reliable. Figure S2 in the Supplements Appendix shows the distribution of total SDQ and of three specific disorders (conduct disorder, emotional disorder, and hyperactivity) at the age of 9. For each scale we show the distribution of the reports of all three informants (parent, teacher, and child self report). In line with the averages reported in table S2 in the Supplements Appendix, the histograms show that teachers generally report lower levels of mental distress compared to parents, and that children report higher levels of mental distress. The same is true for all three disorders and for the total scale. As the histograms suggest, the reports are very different from each other. We conclude that using a measure of mental health based on any single informant would introduce measurement error that might significantly alter the results (see Griliches (28)). Furthermore, we treat the reports as each being a function of a true component and an error component, where the true component is identical for all the reporters. For simplicity, and as it is commonly done in the literature, we assume an additive function. Each reporter thus describes a different aspect of the

child's mental health, and therefore it is harder to assume a mutual component of all the reporters even though the reporters are answering exactly the same questions. It could be that the child behaves differently at home than at school, therefore the reports are different. It may also be the case that the inner personality structure of the child is stable even if she expresses it differently in each environment. Assuming this, the SDQ was initially designed for both teachers and parents. It therefore asks about a wide range of behaviours where some are typical at home and some at school, so that the total score of each report should give the true level of mental distress. We assume that each report contains the true component of the child's mental state. The deviance from this true component is an error term (a formal analysis is presented in the Supplements Appendix). In order to estimate the size of a possible measurement error we look at the correlations between the reports of different informants of same twin's mental health status. Assuming each informant's report has a true and an error component, if the true component is large for all informants then we would expect a high correlation between the reports. Table S3 in the Supplements Appendix shows the correlation matrix between the reports on the twins' SDQ by all three informants (parent, teacher and child self report). In the table, each cell is the correlation between the different mental health variables, where MH_{ji}^{r} is the mental health variable of twin j as reported by informant $r = \{p, t, c\}$, where p, t and c denote parent, teacher and child respectively. The highest between-informants correlation is 0.34 (parent - child self report). This can be interpreted as a measurement error that is 66 per cent of the variation of these variables (see Ashenfelter and Krueger (29)). The correlation matrix also provides information about the existence of an informant-systemic measurement error. Such a measurement error will lead to a higher correlation between $MH9_{1i}^{p}$ and $MH9_{2i}^{p}$ than between $MH9_{1i}^{p}$ and $MH9_{2i}^{t}$ or between $MH9_{1i}^{p}$ and $MH9_{2i}^{c}$. Similarly, such positive correlation will lead

to a higher correlation between $MH9^t_{1i}$ and $MH9^t_{2i}$ than between $MH9^t_{1i}$ and $MH9^p_{2i}$ or between $MH9^t_{1i}$ and $MH9^c_{2i}$, because the same informant's reports contain a common measurement-error component that reports by different informants do not contain. In the presence of classical measurement error these correlations would have been identical. The correlations between the different SDQ reports are consistent with the hypothesis of positively correlated measurement error in the same-person's reports. Given that the psychological literature is not conclusive about who is the best informant, we have made use of all three informants, and used statistical methods in order to minimize the measurement error. In the next sections we outline the main features of our conceptual framework.

Statistical Analysis

The methodological approaches used in this analysis attempt to deal with the biases introduced by the aforementioned measurement error problem and also by omitted variables, including genetic confounders. We start with a simple setup to assess the persistence of mental health,

$$MH12_{1i} = \alpha + \beta MH9_{1i} + \gamma X_i + \mu A_i + u_{1i}$$

$$MH12_{2i} = \alpha + \beta MH9_{2i} + \gamma X_i + \mu A_i + u_{2i}$$

Where $MH12_{1i}, MH12_{2i}$ is our measure of mental health (SDQ) at age 12, and $MH9_{1i}, MH9_{2i}$ is the observed SDQ at age 9 for the first and second twin of the i-th pair respectively. X_i are the observed variables that vary by family, but not between twins. Among these are gender, parental qualification and age of the mother when the first child was born. A_i are the unobservable variables that vary by family, e.g. the family and general environment where the twins were raised, and in case of monozygotic (identical) twins, also all genetic factors. Dizygotic twins on the other hand only share half of the genetic factors. u_{1i}, u_{2i} are the unobservable individual

components. We assume identical equations for the two twins. Our main interest is estimating β, but the OLS estimator may be biased mainly because of two reasons. First, omitted variables: both the family level unobservables, A_i, and the idiosyncratic unobservables, u_{ji} are likely to be correlated with mental health at age 9 and at age 12. Second, measurement error: the complexity of identifying mental health disorders might make the measurement error of $MH9_{ji}$ larger, and the use of within-family data may make this problem worse (see for instance Griliches (28)). Measurement error is also expected in the dependent variable. These issues have been extensively addressed in the economic literature on the returns to schooling also using twin data (28-29). For a more recent critical review, please see Sandewall *et al.* (30).

The size of the omitted variable bias introduced in a least-squares estimation is given by $plim\hat{\beta}_{OLS} = \beta_{OLS} + \mu\frac{Cov(A_i,MH9_{ji})}{Var(MH9_{ji})}$ (mathematical proof is given in Supplements Appendix). It is plausible that the omitted variables, for example A_i, are correlated with the mental health condition at age 9, $\mathrm{MH9}_{ji}$. In such a case, then $Cov(A_i, MH9_{ji}) \neq 0$ and the OLS estimator would be biased. To minimize the bias we would ideally exploit an exogenous shock to mental health at age 9 for $MH9_{ji}$. Obtaining an exogenous variable by generating a mental health shock in an experimental setting is not feasible and is uncommon to obtain in a natural experiment setting. We therefore make the following assumption to identify a proxy for an exogenous mental health shock. Mental health, as any other health problem, is an outcome of the interaction between external factors and internal factors. In the case of a cold for example, the external factor is the density of germs in the air around the person. The internal factor is the quality of the person's immune system. In the case of mental illness the external factor is personal experiences, e.g. exposure to violence or degradation. The internal factor is the personal susceptibility to developing a mental disorder. This assumption is well established in the psychological literature

on mental health, and it is a key concept in modern genetics studies. See, for instance, the concept of genotype-environment (GxE) correlation in Plomin and Asbury (31) and Caspi and Moffitt (32). This assumption leads us to two considerations: (1) The susceptibility to develop a mental health problem (the internal factor) is embedded in the individual's genetics. Obviously, the interaction with the individual's experiences starts at birth, or even before that (33), but the primary location of susceptibility resides in the genes. (2) The source of any difference between two identical twins must therefore come from their experiences, the external factor. Accepting this assumption means therefore that including a twin fixed effect in the regression allows us to control for the internal factor (for identical twins it is 100% of the genes, whereas for non-identical twins the shared part is 50% of the genes) and all experiences the twins have in common. This shared experience is represented by A_i. The remaining difference in mental health after including fixed effects derives from the experience that the twins do not have in common. We refer to this idiosyncratic experience as a shock (not necessarily an exogenous one). Formally, the specification with twin fixed effects (FE) is: $MH12_{ji} = \beta MH9_{ji} + a_i + u_{ji}$. Fixed effects eliminate the unobserved term, A_i and the common components in u_{ji}. FE are asymptotically identical to the first differences (FD) method, which is conducted taking the difference between the two parts of equation: $MH12_{1i} - MH12_{2i} = \beta(MH9_{1i} - MH9_{2i}) + u_{1i} - u_{2i}$. We are using FE rather the FD for efficiency reasons. It could be objected that since the difference in mental health between the twins might stem from experiences back in the womb, or in early childhood, part of the difference in mental health is no more treatable at age 9. We therefore control for some idiosyncratic early experiences. As an example of very early experiences we consider the twins' birth order, as it is associated with APGAR scores at young ages (34). We control for birth order in all our regressions, but its inclusion among the regressors

has a minor effect on our results. In addition, it is possible that spillover effects take place between twins. This will downward bias our FE results since the mental health of one twin is likely to worsen the mental health of the other twin three years later. The results we find are therefore likely to be a lower bound.

Besides omitted variable bias, the estimated coefficient may also be biased due to measurement error. Measurement error (ME) may be a particularly serious issue in our context because of the complexity to identify and correctly report mental states. As shown in Figure S2 in the Supplements Appendix, different informants (parent, teacher and child) provide very different reports on the child's mental health. The psychological literature does not provide clear-cut evidence on which of the informants is more reliable. We therefore use all three reports available in the dataset (parent, teacher and child self report) to minimize the bias. Here we present our empirical framework. In the classical ME model, measured mental health (MH) is related to "true" MH by $MH9_{ji}^{r} = MH9_{ji}^{true} + v_{ji}^{r}$, where $MH9_{ji}^{true}$ is the "true" MH of twin j in family i, and $MH9_{ji}^{r}$ is the measured MH of twin j as reported by informant $r = \{p, t, c\}$ where p, t and c denote parent, teacher and child, respectively. And v_{ji}^{r} are measurement errors that are uncorrelated with $MH9_{ji}^{true}$. If we assume no other bias is in place (no omitted variable bias) the probability limit of the estimate is (mathematical proof is given in Supplements Appendix):

$$plim\ \hat{\beta}_{OLS} = \beta_{OLS}\left[1 - \frac{\mathrm{Var}(v_{ji}^{r})}{\mathrm{Var}(MH9_{ji}^{r})}\right] = \beta_{OLS}\left[1 - \frac{\mathrm{Var}(MH9_{ji}^{true})}{\mathrm{Var}(MH9_{ji}^{r})}\right]$$

where β_{OLS} is the least square coefficient if MH were precisely measured, $\mathrm{Var}(v_{ji}^{r})$ is the presumed variance of the measurement error, $Var\left(MH9_{ji}^{r}\right)$ is the variance of the measured MH at age 9 and $Var\left(MH9_{ji}^{true}\right)$ is the variance of the true MH at age 9. In the previous section we made the case of using FE to deal with the omitted variable problem. By introducing FE,

however, we may inflate the ME problem. Intuitively, while v_{ji}^{r} is purely random ME, and therefore its variance is the same in the population level analysis (OLS) and in the within family analysis (FE), the variance of MH between twins is much smaller than the variance in the population, and therefore the variance of $MH9_{ji}^{r}$ does change when going from OLS to FE. In particular, in FE the nominator of the above equation doesn't change while the denominator is smaller, hence a bigger bias (28). Formally, the probability limit of the within-family estimate is: $plim\ \hat{\beta}_{FE}=\beta_{FE}\left[1-\frac{\mathrm{Var}(v_{ji}^{r})}{\mathrm{Var}(MH9_{ji}^{r})}\cdot\frac{1}{1-\rho_{MH9}}\right]$ where β_{FE} is the FE coefficient if MH were precisely measured, and ρ_{MH9} is the ratio family to total variance components in $MH9_{ji}^{r}$ (the variance of MH9 between two twins inside the family over the variance of all the twins in the sample). In order to minimize the measurement error, we take advantage of the richness of our twin data (three MH reports for each twin) and perform 2SLS, using two reports as instrumental variable for the third one: $MH12_{ji}=\beta(\hat{MH9}_{ji}^{r1})+a_i+u_{ji}$ where $(\hat{MH9}_{ji}^{r1})$ are the predicted values from the regression: $MH9_{ji}^{r1}=\alpha_0+\alpha_1 MH9_{ji}^{r2}+\alpha_2 MH9_{ji}^{r3}+a_i+\varepsilon_i$. We will use one of the following three specifications: $r_1=t, r_2=p, r_3=c$ or $r_1=p, r_2=t, r_3=c$ or $r_1=c, r_2=p, r_3=t$, on the basis of the robustness of the first stage (largest F statistic in the first stage). So far we discussed classical ME. However, this may not be the only type of ME we encounter in our data. The psychological literature suggests that teachers' and parents' responses tend to be biased by the child's learning disabilities and by the child's social class (27). It is therefore plausible to expect that teachers' report to be correlated between them and similar for the parents' reports. Moreover, it is also likely that when parents report the mental health of both their twins, they will provide correlated reports. The correlation matrix (see table S3 in the Supplements Appendix) was in line with such an informant-systemic error. We therefore need to account for the presence of an

informant-systemic error. We consider a ME model given by $MH9_{ji}^{r}$=$MH9_{ji}^{true}$+v_i^r+v_{ji}^r where $MH9_{ji}^{true}$ is the true MH of twin j in family i, $MH9_{ji}^{r}$ is the measured MH status of twin j in family i as reported by informant r_{ptc}, v_i^r is the systemic error of informant r, and v_{ji}^r is the random measurement error component. Both v_i^r and v_{ji}^r are assumed to be uncorrelated with $MH9_{ji}^{true}$. A way to deal with this type of non-classical measurement error is to control for the informant, for instance by differencing two reports of the same informant (e.g. reports on the SDQ of two different twins). Using twin fixed effects does this by eliminating the systemic error component. This is the same strategy we use to deal with the omitted variable bias problem. Another type of non-classical measurement error is the one that is caused by regression towards the mean. Mean reversion takes place when observations who are close to the ends of the distribution can not move away from the mean but only towards it (bounded scales). Mean reversion takes also place when the same variable is observed in two different points in time. If some observations got a score lower (higher) than the mean when first observed, then they have a higher probability to score a higher (lower) value in the second observation. In general, the second observation will be closer to the mean. This will bias our results since this misperception of the true values is affecting more the observations as we move closer to the ends of the distribution. Therefore the coefficient of the regression will be downward biased. For this reason we expect our results to represent the lower bound of the effect. So far we have assumed that our explanatory variable, mental health status at age 9 (MH9) may be measured with error. For the same reasons, it is also plausible to expect that our dependent variable, mental health status at age 12 (MH12) is measured with error. Generally a dependent variable measured with error (but not correlated with the explanatory variable) does not lead to biased estimates. This is, however, not the case here. The measurement error in MH12 is given by $MH12_{ji}^{r} = MH12_{ji}^{true} + e_i^r - e_{ji}^r$

where $MH12_{ji}^{true}$ is the true mental health status (MH) at age 12 of twin j of family i, $MH12_{ji}^{r}$ is the measured MH status at age 12 of twin j of family i as reported by informant r_{ptc}, e_i^r is the systemic error of informant r of family i, and e_{ji}^r is random measurement error. By using FE we control for the informant systemic error. We are therefore left with the random error, e_{ji}^r. To illustrate the bias we expect because of measurement error in MH12, we estimate OLS including the measurement error in both the dependent variable and the explanatory variable ($MH12_{ji}^{r} = MH12_{ji}^{true} + e_{ji}^r$ and $MH9_{ji}^{r} = MH9_{ji}^{true} + v_{ji}^r$). The probability limit is given by (mathematical proof given in Supplements Appendix):

$$p\lim \hat{\beta}_{OLS} = \beta_{OLS}\left[\frac{Var(MH9_{ji}^{true})}{Var(MH9_{ji}^{r})}\right] + \frac{Cov(MH9_{ji}^{true}, e_{ji}^{r}) + Cov(v_{ji}^{r}, e_{ji}^{r})}{Var(MH9_{ji}^{r})}$$

From this equation we see that the estimator is biased if Cov($MH9_{ji}^{true}$, e_{ji}^r) $\neq 0$, and/or if Cov(v_{ji}^r, e_{ji}^r) $\neq 0$. In our case both covariances are plausibly different from zero. In order to minimize this measurement error bias, we should minimize both covariances - Cov($MH9_{ji}^{true}$, e_{ji}^r) and Cov(v_{ji}^r, e_{ji}^r). To do this we choose the dependent variable that has the minimum correlation with the independent variables. Recall that the reduced form we are running is $MH12_{ji}$=γ_0+$\gamma_1 MH9_{ji}^{r_2}$+$\gamma_2 MH9_{ji}^{r_3}$+a_i+v_{ji} (where r_2 and r_3 are any two of the three informants). Therefore, the dependent variable would be $MH12_{ji}^{r_1}$ (where r_1 is the third informant). This is also better (in terms of minimizing correlation with the independent variables) than taking any combination between the three reports at 12 (e.g. their average).

Results

Table 1 reports on the main results. When controlling for twin common traits, including their genetic endowment, we find that 55% of a one standard deviation change in mental health (total SDQ) at age 9 will still be present three years later. When we investigate the SDQ components separately we observe that 64%, 57%, and 63% of a one standard deviation change at age 9 in, respectively, conduct disorder, emotional disorder, and hyperactivity will still be present three years later. We do not interpret the coefficients on peer problems and prosocial behaviour as being unbiased because the analysis is only weakly instrumented (35). The enduring effect of childhood experience on mental health varies by gender. Conduct disorder due to idiosyncratic experience in girls has an enduring effect on later mental health that is double the size as compared to boys. Conduct disorder induced by, for example, a traumatic event would appear to hardly fade in girls over the period of nine to twelve years. The opposite is the case for anxiety and depression where the lasting effects linger more greatly for boys as compared to girls. No material gender-specific differences are obtained for the other subcomponents of the SDQ mental health index. When comparing the results for identical versus non-identical twin pairs we find the effect sizes considerably and consistently higher for non-identical twins (see Table S4 in Supplements Appendix). This upward bias is due to the genetic variation that is unaccounted for in non-identical twin pairs. These comparative statistics imply that genetic variation plays an important role in the development of mental health and that it is imperative to disentangle genetic influences from environmental influences when studying the effect of childhood experience on later mental health. When applying econometric instrumentation to account for measurement error we observe consistent increases in the size of the effects that are similar across identical and non-identical twins (Table S4 in Supplements Appendix). This is the result of accounting for the role of ex ante mental health in shaping childhood experience.

When extending the analysis to include mental health at ages four and seven we find a significant degree of continuity over the considered period of four to twelve years old. This evolution over time is shown in Figure 1 and holds for the SDQ mental health index, hyperactivity and prosocial behaviour, and to a lesser extent for anxiety and depression. Generally, the coefficients become smaller over time indicating a gradually diminishing effect of an earlier idiosyncratic experience on mental health as the child ages (also see Table S5 in Supplements Appendix). Though mental health issues due to experience seem to fade over time a significant residual effect remains. Figure 1 also reveals that the size of the enduring effect of experience on mental health varies by the childhood age at impact. We find that, in general, the older the child is the higher is the continuity of a shock to mental health. For example, considering the overall SDQ mental health score, we observe that the continuity effect between age 4 and 7 is less than the continuity effect between age 7 and 9 which, in turn, is less than the continuity effect between age 9 and 12 (also see Table S5 in Supplements Appendix). A similar pattern is observed for the mental health categories that make up the total SDQ mental health index. This result appears to indicate that child mental health becomes more absorbent over time with experience having an increasingly enduring effect.

Discussion

The results reported here show that there is lasting impact of idiosyncratic experience on mental health in childhood and that its effect diminishes over time. About half of the effect of a personal experience on mental health at age 9 still remains at age 12. Though this report makes an intuitive point it is, to our knowledge, the first unbiased estimate of the enduring effect of childhood experience on mental health. As such, it contributes to the study of mental health and

human capital accumulation in a number of ways. First, our investigation into the continuity of child mental health allows us to get a grasp on the degree to which childhood events may influence later health and socio-economic outcomes by way of their lagged effect on subsequent mental health. Second, we address statistical identification problems not previously dealt with in the mental health literature that allow us to present unbiased results. The important role of genetic variation in mental health (11-16), as well as other unobservable traits such as quality of nutrition in the womb (33), is accounted for using identical twin data. Having multiple independent reports allows us to instrument the analysis to address potential reverse causality (the effect that mental health issues may have on childhood experience) and further reduce measurement error. Third, when considering mental health across gender we observe significant differences in terms of conduct and emotional disorders. Finally, the continuity of mental health and the observed gender differences raise the issue of treatment and its effectiveness. Understanding the evolution of mental health helps identifying when mental health issues can be most effectively treated.

References

1. Layard R et al. How mental illness loses out in the NHS. London (UK): Mental Health Policy Group, Centre for Economic Performance; 2012 June [cited 2012 10 August]. Available from: http://cep.lse.ac.uk/pubs/download/special/cepsp26.pdf

2. Organisation for Economic Co-operation and Development. Mental Health and Work: Evidence, Challenges and Policy Directions. OECD Publishing; 2011.

3. Fergusson D, Horwood J, Ridder E. Show me the child at seven: the consequences of conduct problems in childhood for psychosocial functioning in adulthood. J Child Psychol Psychiatr. 2005 Mar 8;46(8):837-849.

4. Coffey C, Veit F, Wolfe R, et al. Mortality in young offenders: retrospective cohort study. BMJ. 2003 May 17;326(7398):1064.

5. Von Korff M, Scott KM, Gureje O. Global perspectives on mental-physical comorbidity in the WHO World Mental Health Surveys. Cambridge: Cambridge University Press; 2009.

6. Von Korff M, Alonso J, Ormel J, et al. Childhood psychological stressors and adult onset arthritis: Broad spectrum risk factors and allostatic load. Pain. 2009 May;143(1-2):76-83.

7. Walker EA, Gelfrand A, Katon WJ, et al. Adult health status of women with histories of childhood abuse and neglect. Am J Med. 1999 Oct;107(4):332-339.

8. Scott KM, Von Korff M, Angermeyer MC, et al. Association of childhood adversities and early-onset mental disorders with adult-onset chronic physical conditions. Arch Gen Psychiatry. 2011 Aug;68(8):838-844

9. Case A, Fertig A, Paxson C. The lasting impact of childhood health and circumstance. J Health Econ. 2005 Mar;24(2):365–389.

10. Goldsmith AH, Veum JR, Darity W. The impact of psychological and human capital on wages. Econ Inq. 1997 Oct;35(4):815–829.

11. Lewis CM, Levinson DF, Wise LH et al. Genome scan meta-analysis of schizophrenia and bipolar disorder, part II: Schizophrenia. Am J Hum Genet 2003 Jul;73(1):34-48.

12. Sullivan PF, Neale MC, Kendler KS. Genetic epidemiology of major depression: Review and meta-analysis. Am J Psychol. 2000 Oct;157(10):1552-1562.

13. Kendler KS, Prescott CA, Myers J et al. The structure of genetic and environmental risk factors for common psychiatric and substance use disorders in men and women. Arch Gen Psychiatry. 2003 Sep;60(9):929-937.

14. Hettema JM, Neale MC, Kendler KS. A review of and meta-analysis of the genetic epidemiology of anxiety disorders. Am J Psychiatry. 2001 Oct;158(10):1568-1578.

15. Kendler KS. Twin studies of psychiatric illness: An update. Arch Gen Psychiatry. 2001 Nov;58(11):1005-1014.

16. Saudino KJ, Ronald A, Plomin R. The etiology of behaviour problems in 7-year-old twins: substantial genetic influence and negligible shared environmental influence for parent ratings and ratings by same and different teachers. J Abnorm Child Psychol. 2005 Feb;33(1):113-30.

17. Kessler RC, Davis CG, Kendler KS. Childhood adversity and adult psychiatric disorder in the US National Comorbidity Survey. Psychol Med 1997 Sep;27(5):1101-1119.

18. Kendler KS, Bulik CM, Silberg J et al. Childhood sexual and adult psychiatric and substance use disorders in women – An epidemiological and Cotwin control analysis. Arch Gen Psychiatry 2000 Oct;57(10):953-959.

19. Kendler KS, Kuhn JW, Vittum J, Prescott CA, Riley B. The interaction of stressful life events and a serotonin transporter polymorphism in the prediction of episodes of major depression: A replication. Arch Gen Psychiatry 2005 May;62(5):529-535.

20. Rhode P, Lewinsohn PM, Seeley JR. Are adolescents changed by an episode of major depression. J Am Acad Child Adolesc Psychiatry 1994 Nov-Dec;33(9):1289-1298.

21. Rhode P, Lewinsohn PM, Seeley JR. Are people changed by the experience of having an episode of depression – a further test of the scar hypothesis. J Abnorm Psychol 1990 Aug;99(3):264-271.

22. Benjet C, Borges G, Medina-Mora ME. Chronic childhood adversity and onset of psychopathology during three life stages: Childhood, adolescence and adulthood. J Psychiatric Res 2010 Aug;44(11):732-740.

23. Oliver B, Plomin R. Twins Early Development Study (TEDS): A multivariate, longitudinal genetic investigation of language, cognition and behaviour problems from childhood through adolescence. Twin Res and Hum Genet. 2007 Feb 21;10(1):96–105.

24. Goodman R. The strengths and difficulties questionnaire: a research note. J Child Psychol Psychiatr. 1997 Jul;38(5): 581–586.

25. Goodman R. Psychometric properties of the strengths and difficulties questionnaire. J Am Acad of Child Adolesc Psychiatry. 2001 Nov;40(11):1337–1345.

26. Reyes ADL, Kazdin AE. Informant discrepancies in the assessment of childhood psychopathology: A critical review, theoretical framework, and recommendations for further study. Psychol Bull. 2005 Jul;131(4):483–509.

27. Collishaw S, Goodman R, Ford T, et al. How far are associations between child, family and community factors and child psychopathology informant-specific and informant-general? J Child Psychol Psychiatr. 2009 Apr 21;50(5):571–580.

28. Griliches Z. Sibling models and data in economics: Beginnings of a survey. J Polit Econ. 1979 Oct;87(5):S37–S64.

29. Ashenfelter O, Krueger A. Estimates of the economic return to schooling from a new sample of twins. Am Econ Rev. 1994 Dec;84(5): 1157–1173.

30. Sandewall O, Cesarini D, Johannesson M. The co-twin methodology and returns to schooling – testing a critical assumption. IFN Working Paper [Internet]. 2009 Aug 24 [cited 2012 10 August];809. Available from: http://ssrn.com/abstract=1674837

31. Plomin R, Asbury K. Nature and nurture: Genetic and environmental influences on behaviour. Ann Am Acad Pol Soc Sci. 2005 Jul;600(1):86 –98.

32. Caspi A, Moffitt TE. Gene-environment interactions in psychiatry: joining forces with neuroscience. Nat Rev Neurosci. 2006 Jul;7(7):583-590.

33. Barker DJ. The fetal and infant origins of adult disease. BMJ. 1990 Nov 17;301(6761):1111.

34. Young BK et al. Differences in twins: the importance of birth order. Am J Obstet Gynecol. 1985 Apr 1;151(7):915–921.

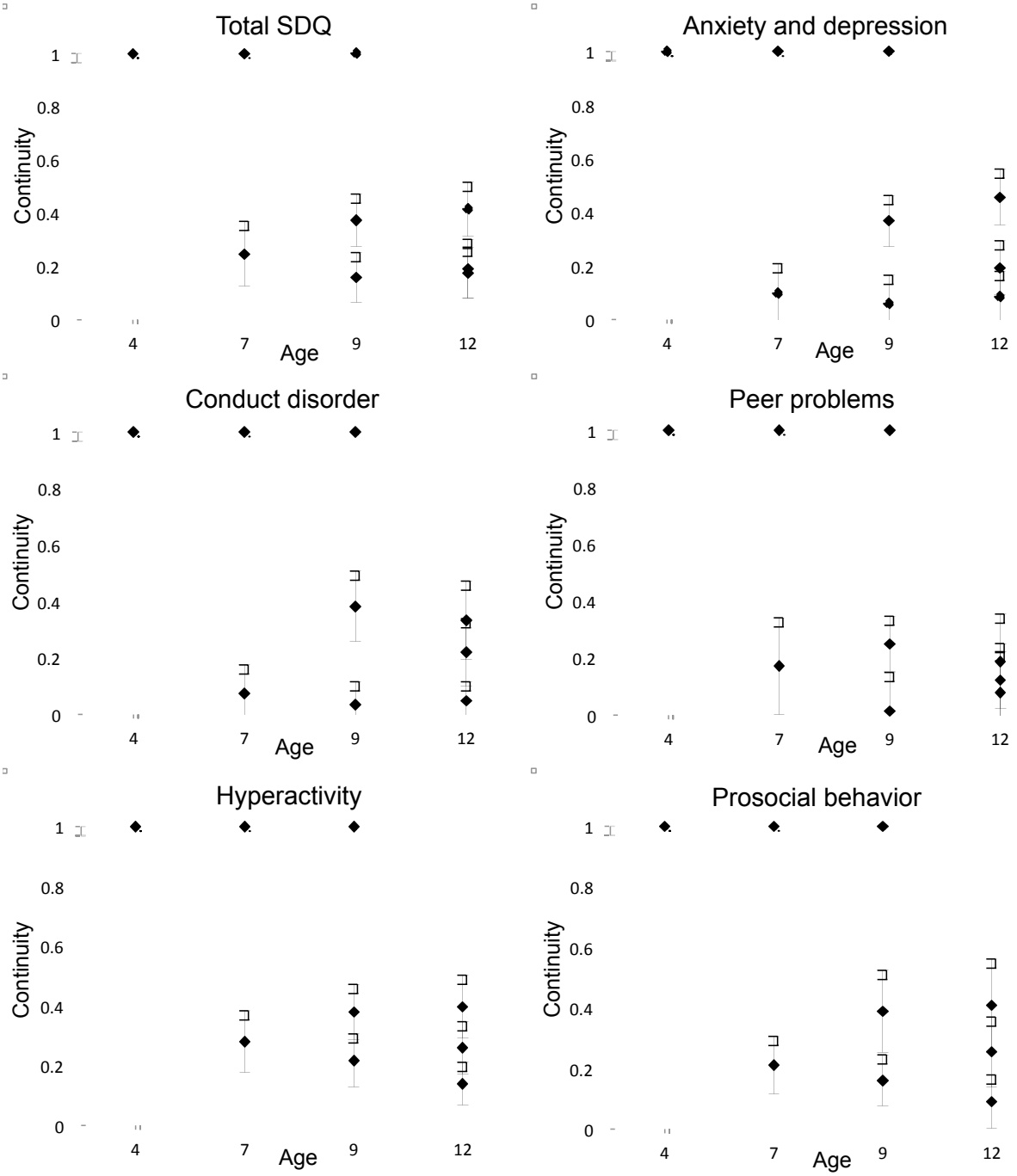
35. Stock JH, Yogo M. Identification and inference for econometric models: essays in honor of Thomas Rothenberg. Cambridge: Cambridge Univ. Press; c2005. 5, Testing for Weak Instruments in Linear IV Regression; p. 80-108.

Table 1. An instrumented co-twin analysis of the effect of experience at age 9 on mental health age 12

	Entire MZ sample (1)		MZ Boys (2)		MZ Girls (3)	
Total SDQ	0.552	**	0.598	**	0.501	**
SE	(0.102)		(0.163)		(0.124)	
N	1350		562		788	
first stage F-statistic	35.3		17.6		17.9	
Conduct disorder	0.635	**	0.404	**	0.933	**
SE	(0.135)		(0.164)		(0.243)	
N	1330		550		780	
first stage F-statistic	25.3		18.7		8.7	
Anxiety and depression	0.567	**	0.712	**	0.476	**
SE	(0.129)		(0.179)		(0.175)	
N	1344		560		784	
first stage F-statistic	29.2		12.4		17.1	
Hyperactivity	0.626	**	0.668	**	0.586	**
SE	(0.100)		(0.154)		(0.126)	
N	1328		550		778	
first stage F-statistic	38.6		19.1		20.3	
Peer problems	0.505	**	0.848	**	0.207	
SE	(0.205)		(0.411)		(0.182)	
N	1340		554		786	
first stage F-statistic	8.0		3.1		4.9	
Prosocial behaviour	0.345		0.186		0.747	
SE	(0.218)		(0.226)		(0.544)	
N	1342		562		780	
first stage F-statistic	6.9		6.5		1.4	

Note: Shown are standardized coefficients for the persistence of differences in mental health between identical or monozygotic (MZ) twins (due to idiosyncratic experience). Mental health is measured by the Strengths and Difficulties Questionnaire (SDQ), and its sub-indices, and is evaluated by the parents. The child self-report and teacher report are used to instrument the analysis and reduce measurement error. Coefficients are also reported for the sample split by gender. All regressions control for birth order. Robust SEs are in parenthesis. One star represents 90% significance and two stars represent 95% significance. The number of observations and the first stage F-statistic are reported underneath the SEs.

Fig. 1. The effect of experience at ages 4, 7, and 9 on later mental health.



Note: Shown is the persistence of differences in mental health between identical twins (due to idiosyncratic experience). Mental health is measured by the Strengths and Difficulties Questionnaire (SDQ), and its sub-indices, and is evaluated by the parents.

Supplements Appendix:

The Enduring Impact of Childhood Experience on Mental Health: Evidence Using Instrumented Co-Twin Data

Rachel Berner Shalem, Francesca Cornaglia, Jan-Emmanuel De Neve

Figures S1-S2
Tables S1-S5
Mathematical proofs

Figure S1. The distribution of total SDQ (parent report) at age 9 and at 12 in the TEDS sample.

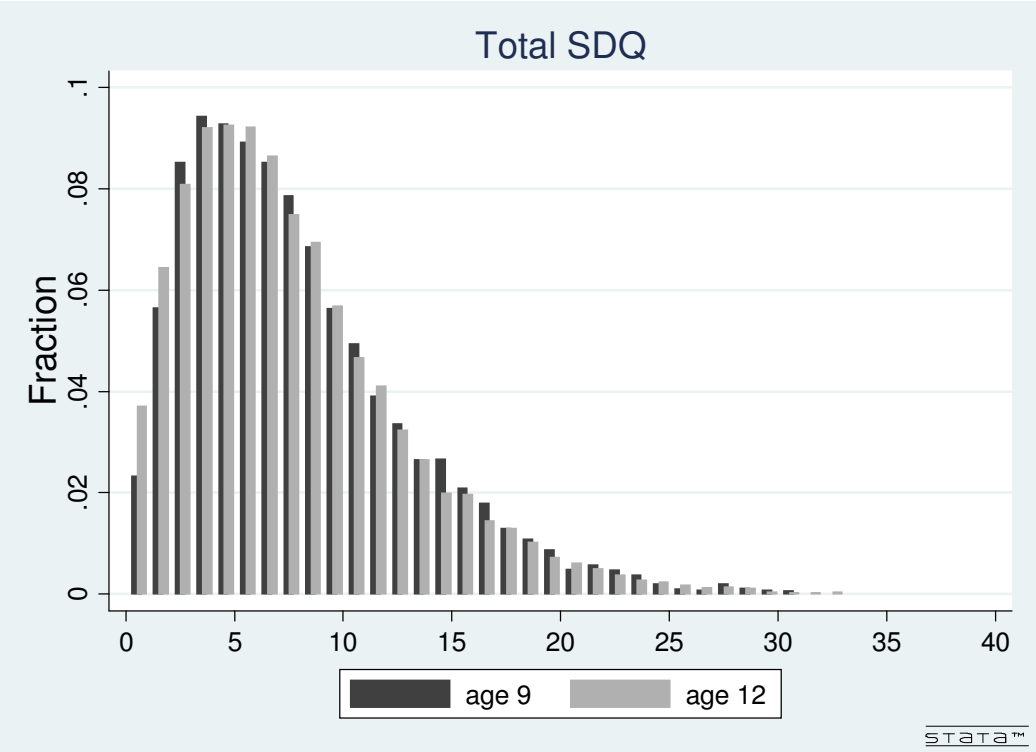
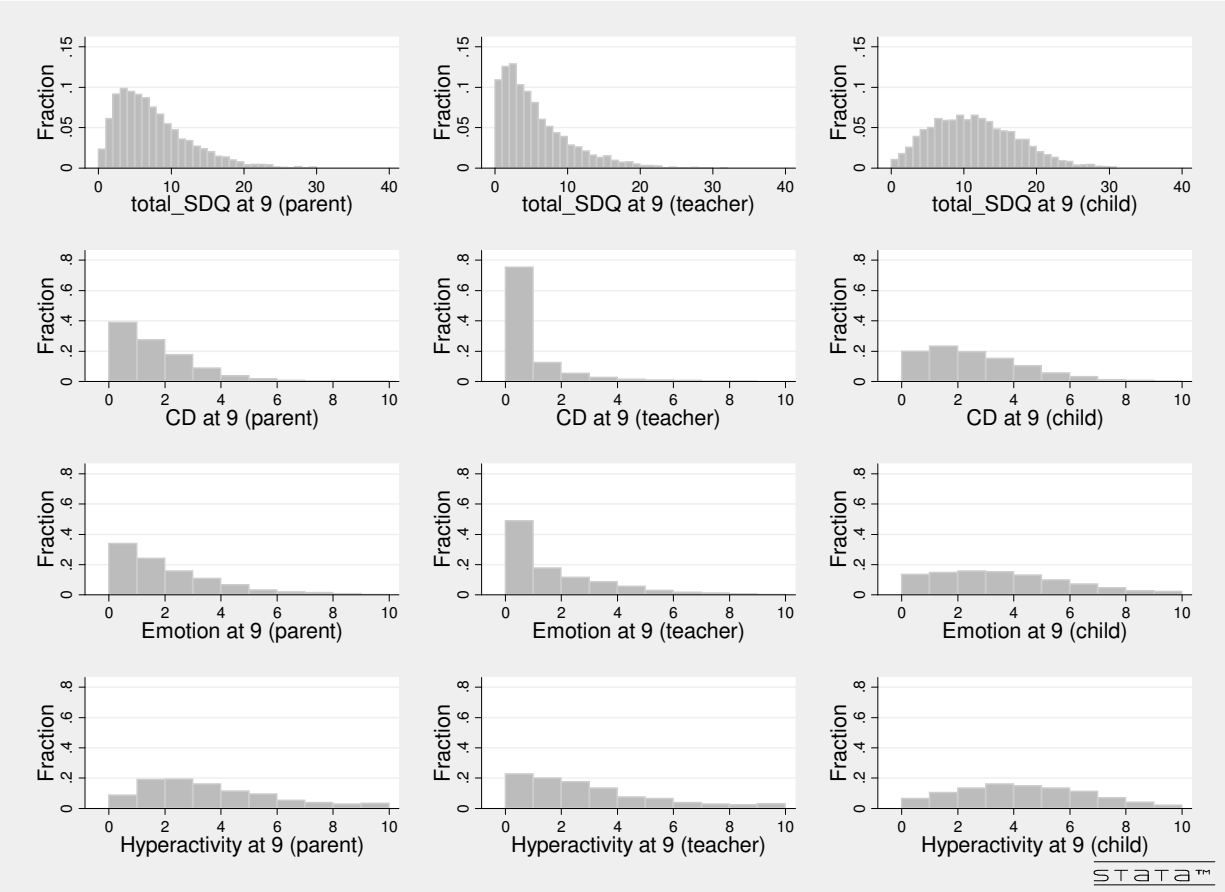


Figure S2. SDQ distributions by informant at age 9.



Note: The histograms show the distribution of Total SDQ (top row) and of three specific disorders: conduct disorder (second row), emotional disorder (third row), and hyperactivity (bottom row). All graphs refer to age 9. For each mental health measure we show the distribution of the reports for all three informants: parent SDQ on the left, teacher SDQ in the middle and child self report in the right column.

Table S1. The Strengths and Difficulties Questionnaire.

Emotional Symptoms Scale	Not True	Somewhat True	Certainly True
Often complains of headaches, stomach-aches or sickness	0	1	2
Many worries, often seems worried	0	1	2
Often unhappy, down-hearted or tearful	0	1	2
Nervous or clingy in new situations, easily loses confidence	0	1	2
Many fears, easily scared	0	1	2

Conduct disorder	Not True	Somewhat True	Certainly True
Often has temper tantrums or hot tempers	0	1	2
Often fights with other children or bullies them	0	1	2
Generally obedient, usually does what adults request	0	1	2
Often lies or cheats	0	1	2
Steals from home, school or elsewhere	0	1	2

Hyperactivity	Not True	Somewhat True	Certainly True
Restless, overactive, cannot stay still for long	0	1	2
Constantly fidgeting or squirming	0	1	2
Easily distracted, concentration wanders	0	1	2
Thinks things out before acting	0	1	2
Sees tasks through to the end, good attention span	0	1	2

Peer problems	Not True	Somewhat True	Certainly True
Rather solitary, tends to play alone	0	1	2
Has at least one good friend	0	1	2
Generally liked by other children	0	1	2
Picked on or bullied by other children	0	1	2
Gets on better with adults than with other children	0	1	2

Prosocial behaviour	Not True	Somewhat True	Certainly True
Considerate of other people's feelings	0	1	2
Shares readily with other children	0	1	2
Helpful if someone is hurt, upset or feeling ill	0	1	2
Kind to younger children	0	1	2
Often volunteers to help others	0	1	2

Note: Summing the scores from all the scales except the prosocial scale generates the SDQ score.

Table S2. Mean SDQ score for TEDS Sample and UK population.

panel	Scale	TEDS age 9 (1)	Populatio age 5-10 (2)	signi. (3)	TEDS age 12 (4)	Populatio age 11- (5)	signi. (6)	TEDS age 12 (7)	TEDS age 12 (8)	signi. (9)
	Parent SDQ	N=3596	N=5855		N=3596	N=4443		N=1617	N=1979	
	Total	6.7 (4.8)	8.6 (5.7)	**	6.3 (4.8)	8.2 (5.8)	**	7.0 (4.9)	5.7 (4.5)	**
	Emotional Disorder	1.6 (1.8)	1.9 (2.0)	**	1.6 (1.8)	1.9 (2.0)	**	1.5 (1.7)	1.7 (1.9)	**
A	Conduct Disorder	1.2 (1.3)	1.6 (1.7)	**	1.2 (1.4)	1.5 (1.7)	**	1.3 (1.4)	1.1 (1.3)	**
	Hyperactivity	3.0 (2.3)	3.6 (2.7)	**	2.5 (2.1)	3.2 (2.6)	**	3.1 (2.3)	2.0 (1.8)	**
	Peer Problems	0.9 (1.4)	1.4 (1.7)	**	1.0 (1.4)	1.5 (1.7)	**	1.1 (1.6)	0.8 (1.3)	**
	Prosocial Behaviour	8.3 (1.7)	8.6 (1.6)	**	8.6 (1.6)	8.6 (1.6)		8.3 (1.7)	8.9 (1.4)	**
	Teacher SDQ	N=3596	N=4801		N=2481	N=3407		N=1084	N=1397	
	Total	5.0 (4.6)	6.7 (5.9)	**	4.6 (4.7)	6.3 (6.1)	**	5.5 (5.1)	3.8 (4.2)	**
	Emotional Disorder	1.3 (1.8)	1.5 (1.9)	**	1.1 (1.7)	1.3 (1.9)	**	1.1 (1.7)	1.1 (1.7)	
B	Conduct Disorder	0.5 (1.1)	0.9 (1.6)	**	0.4 (1.1)	0.9 (1.7)	**	0.6 (1.3)	0.3 (.9)	**
	Hyperactivity	2.4 (2.4)	3.0 (2.8)	**	2.0 (2.3)	2.6 (2.7)	**	2.6 (2.7)	1.5 (1.9)	**
	Peer Problems	0.8 (1.4)	1.4 (1.8)	**	1.1 (1.6)	1.4 (1.8)	**	1.3 (1.7)	1.0 (1.5)	**
	Prosocial Behaviour	7.6 (2.2)	7.3 (2.4)	**	8.0 (2.0)	7.1 (2.4)	**	7.4 (2.1)	8.4 (1.8)	**
	Self-report SDQ	N=3596			N=3552	N=4228		N=1597	N=1955	
	Total	10.9 (5.8)	N/A		8.5 (5.2)	10.3 (5.2)	**	9.0 (5.4)	8.0 (5.1)	**
	Emotional Disorder	3.2 (2.3)	N/A		2.1 (2.0)	2.8 (2.1)	**	1.8 (1.9)	2.3 (2.1)	**
C	Conduct Disorder	2.1 (1.8)	N/A		1.8 (1.6)	2.2 (1.7)	**	2.0 (1.7)	1.6 (1.5)	**
	Hyperactivity	3.8 (2.3)	N/A		3.4 (2.3)	3.8 (2.2)	**	3.8 (2.4)	3.0 (2.1)	**
	Peer Problems	1.8 (1.7)	N/A		1.2 (1.5)	1.5 (1.4)	**	1.4 (1.6)	1.1 (1.4)	**
	Prosocial Behaviour	7.9 (1.8)	N/A		7.3 (1.9)	8.0 (1.7)	**	6.7 (1.9)	7.8 (1.7)	**

Note: Columns (1) and (4) refer to the TEDS data, age 9 and 12 respectively. Columns (2) and (5) refer to the UK population, age 9 and 12 respectively (Source: National Survey of Child and Adolescent Mental Health (1999) in http://www.sdqinfo.org/norms/UKNorm3.pdf). Columns (7) and (8) refer to the TEDS data, age 12, boys and girls' subsamples respectively. Columns (3), (6) and (9) show whether the means of the two prior columns (7) and (8) are significantly different from each other. Two stars indicate 95% significance level. Panel A, B and C are the child's SDQ level as reported by the parent, teacher, and child self-report. Each panel shows the total SDQ level, four disorder specific scales and the prosocial behaviour scale. Standard deviation is in parenthesis.

Table S3. Pearson correlation matrix of SDQ scores for identical twins.

	parent on 1st twin ($MH9^{p}{}_{1}$)	teacher on 1st twin ($MH9^{t}{}_{1}$)	1st twin - self report ($MH9^{c}{}_{1}$)	parent on 2st twin ($MH9^{p}{}_{2}$)	teacher on 2st twin ($MH9^{t}{}_{2}$)	2st twin - self report ($MH9^{c}{}_{2}$)
parent on 1st twin ($MH9^{p}{}_{1}$)	1.000					
teacher on 1st twin ($MH9^{t}{}_{1}$)	0.302	1.000				
1st twin - self report ($MH9^{c}{}_{1}$)	0.341	0.234	1.000			
parent on 2st twin ($MH9^{p}{}_{2}$)	0.773	0.290	0.238	1.000		
teacher on 2st twin ($MH9^{t}{}_{2}$)	0.205	0.595	0.155	0.231	1.000	
2st twin - self report ($MH9^{c}{}_{2}$)	0.240	0.115	0.463	0.329	0.104	1.000

Table S4. The effect of experience at age 9 on mental health at age 12 using different model specifications

	OLS (1)	FE (2)	2SLS (3)	FE + 2SLS (4)
Panel A - Non Identical Twins				
Total SDQ	0.693 **	0.599 **	0.881 **	0.724 **
SE	(0.021)	(0.031)	(0.036)	(0.051)
N	2246	2246	2246	2246
first stage F-statistic	---	---	382.2	194.5
Conduct disorder	0.586 **	0.556 **	0.782 **	0.777 **
SE	(0.024)	(0.039)	(0.045)	(0.067)
N	2234	2234	2234	2234
first stage F-statistic	---	---	197.2	91.4
Anxiety and depression	0.535 **	0.484 **	0.791 **	0.656 **
SE	(0.024)	(0.037)	(0.055)	(0.082)
N	2240	2240	2240	2240
first stage F-statistic	---	---	140.8	59.5
Hyperactivity	0.676 **	0.682 **	0.909 **	0.820 **
SE	(0.018)	(0.024)	(0.032)	(0.039)
N	2226	2226	2226	2226
first stage F-statistic	---	---	376.1	241.7
Peer problems	0.515 **	0.468 **	0.811 **	0.723 **
SE	(0.027)	(0.035)	(0.051)	(0.074)
N	2224	2224	2224	2224
first stage F-statistic	---	---	178.8	90.1
Prosocial behaviour	0.523 **	0.530 **	0.751 **	0.742 **
SE	(0.021)	(0.031)	(0.048)	(0.067)
N	2252	2252	2252	2252
first stage F-statistic	---	---	196.9	72.2

	OLS (1)	FE (2)	2SLS (3)	FE + 2SLS (4)
Panel B - Identical Twins				
Total SDQ	0.631 **	0.391 **	0.855 **	0.552 **
SE	(0.025)	(0.044)	(0.042)	(0.102)
N	1350	1350	1350	1350
first stage F-statistic	---	---	214.6	35.3
Conduct disorder	0.532 **	0.314 **	0.850 **	0.635 **
SE	(0.030)	(0.057)	(0.076)	(0.135)
N	1330	1330	1330	1330
first stage F-statistic	---	---	73.6	25.3
Anxiety and depression	0.502 **	0.371 **	0.767 **	0.567 **
SE	(0.028)	(0.047)	(0.063)	(0.129)
N	1344	1344	1344	1344
first stage F-statistic	---	---	104.8	29.2
Hyperactivity	0.629 **	0.394 **	0.930 **	0.626 **
SE	(0.022)	(0.049)	(0.042)	(0.100)
N	1328	1328	1328	1328
first stage F-statistic	---	---	219.7	38.6
Peer problems	0.477 **	0.205 **	0.854 **	0.505 **
SE	(0.044)	(0.053)	(0.085)	(0.205)
N	1340	1340	1340	1340
first stage F-statistic	---	---	78.1	8.0
Prosocial behaviour	0.531 **	0.410 **	0.817 **	0.345
SE	(0.029)	(0.067)	(0.076)	(0.218)
N	1342	1342	1342	1342
first stage F-statistic	---	---	86.0	6.9
Identification problems taken care of:				
Family OVB (for MZ twins, including genes)		✓		✓
ME in explanatory variable			✓	✓
ME in dependent variable			✓	✓
Controlling birth order	✓	✓	✓	✓

Note: This table shows the standardized regression coefficients for the SDQ measures at age 9 and their effect on lagged SDQ measures (age 12). Panel A and B show results for non-identical and identical twins, respectively. All regressions control for birth order. Robust SEs are in parenthesis. Two stars indicate 95% significance. Number of observations are reported underneath the SE and first stage F-statistic is reported for the instrumented models in columns (3) and (4).

Table S5. Effects of experience on mental health between ages 4, 7, 9 and 12 in identical twin data.

Identical Twins		9 to 12 (1)	7 to 12 (2)	4 to 12 (3)	7 to 9 (4)	4 to 9 (5)	4 to 7 (6)
Total SDQ		0.415 **	0.191 **	0.176 **	0.373 **	0.159 **	0.247 **
	SE	(0.050)	(0.055)	(0.048)	(0.049)	(0.047)	(0.061)
	N	926	926	926	926	926	926
Conduct disorder		0.334 **	0.220 **	0.049	0.383 **	0.034	0.074
	SE	(0.069)	(0.059)	(0.033)	(0.062)	(0.041)	(0.050)
	N	906	906	906	906	906	906
Anxiety and depression		0.457 **	0.193 **	0.084 *	0.368 **	0.059	0.097 *
	SE	(0.052)	(0.050)	(0.047)	(0.047)	(0.053)	(0.056)
	N	920	920	920	920	920	920
Hyperactivity		0.398 **	0.260 **	0.140 **	0.380 **	0.217 **	0.280 **
	SE	(0.052)	(0.043)	(0.036)	(0.046)	(0.045)	(0.052)
	N	902	902	902	902	902	902
Peer problems		0.188 **	0.123 *	0.079	0.251 **	0.014	0.172 **
	SE	(0.083)	(0.064)	(0.072)	(0.048)	(0.068)	(0.085)
	N	922	922	922	922	922	922
Prosocial behaviour		0.408 **	0.254 **	0.090 **	0.389 **	0.160 **	0.211 **
	SE	(0.077)	(0.058)	(0.044)	(0.068)	(0.042)	(0.047)
	N	920	920	920	920	920	920

Note: Each column reports the standardized regression coefficients of the persistence of experience on mental health between different age intervals. Figure 1 is a visual representation of these results. Note that these results are obtained using within identical twin variation but are not instrumented because access to all three independent reports on child mental health is only available between ages 9 and 12.

Mathematical proofs.

1. Probability limit of the estimate when omitted variables are present (controls are taken off of the model to simplify the demonstration of the effect of the omitted variables):

$$plim\ \hat{\hat{\beta}}_{OLS} = \frac{Cov\left[MH9_{ji}, MH12_{ji}\right]}{Var\left(MH9_{ji}\right)} = \frac{Cov\left[MH9_{ji}, \left(\alpha + \beta MH9_{ji} + \mu A_i + u_{ji}\right)\right]}{Var\left(MH9_{ji}\right)}$$

$$= \frac{Cov\left(MH9_{ji}, \alpha\right) + Cov\left(MH9_{ji}, \beta MH9_{ji}\right) + Cov\left(MH9_{ji}, \mu A_i\right) + Cov\left(MH9_{ji}, u_{ji}\right)}{Var\left(MH9_{ji}\right)} =$$

$$= \frac{\beta Var\left(MH9_{ji}\right) + \mu Cov\left(MH9_{ji}, A_i\right)}{Var\left(MH9_{ji}\right)} = \beta + \mu \frac{Cov(A_{ji}, MH9_{ji})}{Var(MH9_{ji})}$$

2. Probability limit of the estimate in case of measurement error in the explanatory, $MH9^r_{ji}=MH9_{ji}+v^r_{ji}$ (controls and unobservable variables are taken off to simplify the demonstration of the effect of measurement error):

$$plim\ \hat{\beta}_{OLS} = \frac{Cov\left[MH9^r_{ji}, MH12_{ji}\right]}{Var\left(MH9^r_{ji}\right)} = \frac{Cov\left[MH9^r_{ji}, \left(\alpha + \beta MH9_{ji} + u_{ji}\right)\right]}{Var\left(MH9^r_{ji}\right)}$$

$$= \frac{Cov\left[MH9^r_{ji}, \left(\alpha + \beta MH9^r_{ji} - \beta v^r_{ji} + u_{ji}\right)\right]}{Var\left(MH9^r_{ji}\right)} =$$

$$= \frac{Cov\left(MH9^r_{ji}, \alpha\right) + Cov\left(MH9^r_{ji}, \beta MH9^r_{ji}\right) + Cov\left(MH9^r_{ji}, \left(-\beta v^r_{ji}\right)\right) + Cov\left(MH9^r_{ji}, u_{ji}\right)}{Var\left(MH9^r_{ji}\right)} =$$

$$= \frac{\beta Var\left(MH9^r_{ji}\right) - \beta Cov\left(MH9^r_{ji}, v^r_{ji}\right)}{Var\left(MH9^r_{ji}\right)} = \frac{\beta Var\left(MH9^r_{ji}\right) - \beta Cov\left(\left(MH9^r_{ji} + v^r_{ji}\right), v^r_{ij}\right)}{Var\left(MH9^r_{ji}\right)} =$$

$$= \frac{\beta Var\left(MH9^r_{ji}\right) - \beta Cov\left(MH9_{ji}, v^r_{ji}\right) - \beta Var\left(v^r_{ji}\right)}{Var\left(MH9^r_{ji}\right)}$$

Assuming that the error term, v_{ji}^r is uncorrelated with the true MH9:

$$\frac{\beta Var\left(MH9_{ji}^r\right) - \beta Cov\left(MH9_{ji}, v_{ji}^r\right) - \beta Var\left(v_{ji}^r\right)}{Var\left(MH9_{ji}^r\right)} = \beta\left[1 - \frac{Var\left(v_{ji}^r\right)}{Var\left(MH9_{ji}^r\right)}\right]$$

$$= \beta_{OLS}\left[\frac{Var\left(MH9_{ji}\right)}{Var\left(MH9_{ji}^r\right)}\right]$$

3. Probability limit of the estimate in case of measurement error in the explanatory variable, $MH9_{ji}^r = MH9_{ji} + v_{ji}^r$, and a measurement error in the dependent variable, $MH12_{ji}^r$=$MH12_{ji}$+e_{ji}^r

$$plim\ \hat{\beta}_{OLS} = \frac{Cov\left[MH9_{ji}^r, MH12_{ji}\right]}{Var\left(MH9_{ji}^r\right)} = \frac{Cov\left[MH9_{ji}^r, \left(\alpha + \beta MH9_{ji} + u_{ji}\right)\right]}{Var\left(MH9_{ji}^r\right)}$$

$$= \frac{Cov\left[MH9_{ji}^r, \left(\alpha + \beta MH9_{ji}^r - \beta v_{ji}^r + u_{ji}\right)\right]}{Var\left(MH9_{ji}^r\right)} =$$

$$= \frac{\beta Var\left(MH9_{ji}^r\right) - \beta Cov\left(MH9_{ji}^r, v_{ji}^r\right) + Cov\left(MH9_{ji}^r, e_{ji}^r\right)}{Var\left(MH9_{ji}^r\right)} =$$

$$= \frac{\beta Var\left(MH9_{ji}^r\right) - \beta Cov\left(MH9_{ji}, v_{ji}^r\right) - \beta Var\left(v_{ji}^r\right) + Cov\left(MH9_{ji}, e_{ji}^r\right) + Cov\left(v_{ji}^r, e_{ji}^r\right)}{Var\left(MH9_{ji}^r\right)} =$$

$$= \beta_{OLS}\left[\frac{Var\left(MH9_{ji}\right)}{Var\left(MH9_{ji}^r\right)}\right] + \frac{Cov\left(MH9_{ji}, e_{ji}^r\right) + Cov\left(v_{ji}^r, e_{ji}^r\right)}{Var\left(MH9_{ji}^r\right)}$$

CENTRE FOR ECONOMIC PERFORMANCE
Recent Discussion Papers

1174	Monika Mrázová J. Peter Neary	Selection Effects with Heterogeneous Firms
1173	Nattavudh Powdthavee	Resilience to Economic Shocks and the Long Reach of Childhood Bullying
1172	Gianluca Benigno Huigang Chen Christopher Otrok Alessandro Rebucci Eric R. Young	Optimal Policy for Macro-Financial Stability
1171	Ana Damas de Matos	The Careers of Immigrants
1170	Bianca De Paoli Pawel Zabczyk	Policy Design in a Model with Swings in Risk Appetite
1169	Mirabelle Muûls	Exporters, Importers and Credit Constraints
1168	Thomas Sampson	Brain Drain or Brain Gain? Technology Diffusion and Learning On-the-job
1167	Jérôme Adda	Taxes, Cigarette Consumption, and Smoking Intensity: Reply
1166	Jonathan Wadsworth	Musn't Grumble. Immigration, Health and Health Service Use in the UK and Germany
1165	Nattavudh Powdthavee James Vernoit	The Transferable Scars: A Longitudinal Evidence of Psychological Impact of Past Parental Unemployment on Adolescents in the United Kingdom
1164	Natalie Chen Dennis Novy	On the Measurement of Trade Costs: Direct vs. Indirect Approaches to Quantifying Standards and Technical Regulations
1163	Jörn-Stephan Pischke Hannes Schwandt	A Cautionary Note on Using Industry Affiliation to Predict Income
1162	Cletus C. Coughlin Dennis Novy	Is the International Border Effect Larger than the Domestic Border Effect? Evidence from U.S. Trade
1161	Gianluca Benigno Luca Fornaro	Reserve Accumulation, Growth and Financial Crises

1160	Gianluca Benigno Huigang Chen Christopher Otrok Alessandro Rebucci Eric R. Young	Capital Controls or Exchange Rate Policy? A Pecuniary Externality Perspective
1159	Paul Dolan Georgios Kavetsos	Happy Talk: Mode of Administration Effects on Subjective Well-Being
1158	Alan Manning	Steady-State Equilibrium in a Model of Short-Term Wage-Posting
1157	Joan Costa-Font Mireia Jofre-Bonet Steven T. Yen	Not all Incentives Wash out the Warm Glow: The Case of Blood Donation Revisited
1156	Christian Siegel	Female Employment and Fertility - The Effects of Rising Female Wages
1155	Albrecht Ritschl	The German Transfer Problem, 1920-1933: A Sovereign Debt Perspective
1154	Gabriel M. Ahlfeldt Stephen J. Redding Daniel M. Sturm Nikolaus Wolf	The Economics of Density: Evidence from the Berlin Wall
1153	Nattavudh Powdthavee Yohanes E. Riyanto	Why Do People Pay for Useless Advice?
1152	Thomas Sampson	Selection into Trade and Wage Inequality
1151	Tim Barmby Alex Bryson Barbara Eberth	Human Capital, Matching and Job Satisfaction
1150	Ralf Martin Mirabelle Muûls Laure de Preux Ulrich J. Wagner	Industry Compensation Under Relocation Risk: A Firm-Level Analysis of the EU Emissions Trading Scheme
1149	Albrecht Ritschl	Reparations, Deficits, and Debt Default: the Great Depression in Germany
1148	Alex Bryson John Forth Minghai Zhou	The CEO Labour Market in China's Public Listed Companies

The Centre for Economic Performance Publications Unit
Tel 020 7955 7673 Fax 020 7955 7595
Email info@cep.lse.ac.uk Web site http://cep.lse.ac.uk